CS130 Assignment #4
Date assigned: Tuesday, March 8, 2011 **Date due:** Tuesday, March 16, 2011 4:30 pm
Points: 50

Goals:PASW Regression, PASW Hypothesis testing, Importing Data

Create a Word document **WordAnswers4PUNuetID.docx** using your PUNetID that will contain answers to each of the following questions in order. The Word document and all other files created are to be placed in a folder called **PUNetIDSolution4** once again using your PUNetID.

Answer each question in your Word document under the header **Problem# - Question #.**

Problem #1 Use PASW to solve this problem

Let's revisit the California Economic Indicators data from the previous assignment. Download the same spreadsheet and import the worksheet **CEI3 Leading Indicators** into PASW. Be sure to set the type and measure for each variable correctly and add values and/or Recode the data as we did in class with the Weather Underground data. It may be best to copy and paste this worksheet into a new Excel file before importing it into PASW.

Save this file as **CEI_PUNet.sav.** Answer the following questions in the Word document.

►Use correct spelling, punctuation, grammar, and capitalization in your answers. Label and format all charts professionally.

1) Run a linear, quadratic (x^2), and exponential regression to determine how well **Unemployment Insurance Claims** predicts **Average Weekly Hours**. Copy the resulting chart into the Word document. Copy the Model Summary and Parameter Estimate table to the Word Document. Write (type) the equation for each regression into the Word document, properly labeled, along with that equation's R^2 value.

2) Which regression gave the best model for the data? What is that equation's R^2 value?

3) What number of Average Weekly Hours is predicted by each of the above models (regressions) when Unemployment Insurance Claims is 70,000?

4) What number of Unemployment Insurance Claims is predicted by the linear regression model when Average Weekly Hour is 40.7?

5) What number of Unemployment Insurance Claims is predicted by the exponential regression model when Overtime Hours is 39?

6) Build a table to show the Mean and Standard Deviation for Unemployment Claims broken down by month.

7) Build a Bar Chart to show the mean Unemployment Insurance Claims per month. Make sure the bar in the chart display the months in order, Jan to Dec.

8) Build a Pie Chart to show the sum of the Average Weekly Hours per month.

Problem #2 Use PASW to solve this problem.

Import (any way you choose) the following data into PASW. Name the file PUNetID_Assign4Prob2.sav.

ID	Gender	Height (meters)	Average Steps per Day	Average Calories Burned Per Day
1	f	1.61	8789	2930
2	f	1.44	5072	1652
3	f	1.75	4297	1360
4	f	1.62	7570	2431
5	f	1.59	8997	2980
6	f	1.57	5780	1881
7	f	1.71	6475	2205
8	f	1.56	6646	2194
9	f	1.64	5437	1822
10	f	1.81	5458	1810
11	f	1.59	4630	1568
12	m	2.03	9375	3054
13	m	1.68	4802	1657
14	m	1.91	5974	1914
15	m	1.88	9479	3161
16	m	1.65	9576	3100
17	m	1.96	6654	2139
18	m	2.00	4791	1676
19	m	1.98	7159	2425
20	m	1.92	9812	3192

►Use correct spelling, punctuation, grammar, and capitalization in your answers. Label and format all charts professionally. In the Word document described above, answer each of the following questions in order under the heading specified. That is, list the heading in bold and then give your answer.

►You are to find if the Average Calories Burned Per Day by women is significantly different than the Average Calories Burned Per Day by men.

1) What hypothesis test will you need to perform to find this result? Fully explain why you selected the test you did.
2) State the NULL Hypothesis for the statistical test you selected.
3) Paste in the results of the statistical test.
4) State your conclusion.

5) Explain exactly what the Sig. (2-tailed) value means in this case. I don't want just if the value is less than 0.05 reject the null hypothesis or greater than 0.05 accept the null hypothesis, but in terms of the problem explain the meaningof the Sig. (2-tailed) value including the concept of chance occurrences. Someone without a statistics background must be able to understand your explanation.

► Does the mean height of all the subjects differ from 1.77 meters in a statistically significant way?

6) What hypothesis test will you need to perform to find this result? Fully explain why you selected the test you did.

7) State the NULL Hypothesis for the statistical test you selected.

8) Paste in the results of the statistical test.

9) State your conclusion.

10) How well does Average Steps per Day predict Average Calories Burned per Day? Back up your answer with data.

Submitting your work:

To submit your work, copy your single folder **PUNetIDSolution4** (for me that's will4614Solution4) containing (**WordAnswers4PUNuetID.docx**, **CEI_PUNetID.sav, PUNetID_Assign4Prob2.sav**) into the **CS130 Drop** folder on Turing.

Grading:

1) Correctness of your results
2) Completeness of your results
3) Professional look and correctness of your Word document answers.

Be sure to come see me early with any questions! Also, make sure you reread the Academic Dishonesty policy from the course syllabus.

This is NOT a group project.